A Systems Approach for Governable & Sustainable AI
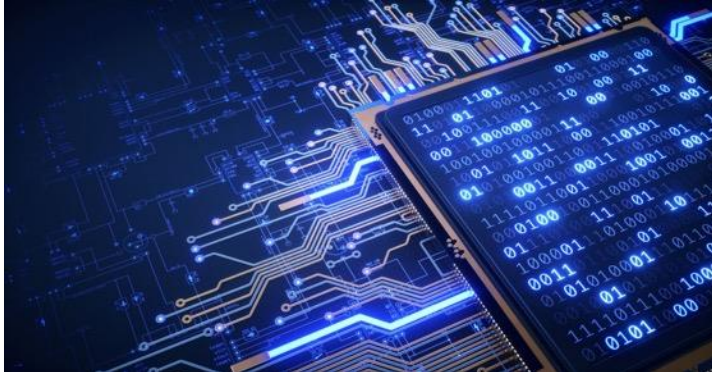
AI4SE & SE4AI
RESEARCH AND APPLICATION WORKSHOP
📅 September 17-18, 2025
📍 Washington, DC

SE4AI 2025

Phil DiBona, LM Fellow, AI Integration
Seth Merrifield, AI Machine Learning Engineer Staff
Lockheed Martin AI Center

LOCKHEED MARTIN

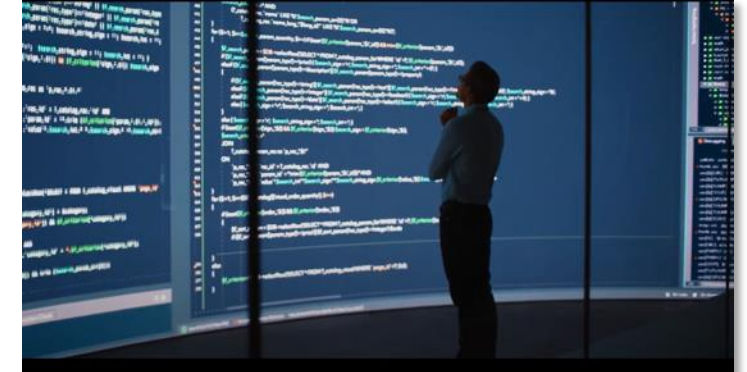Lockheed Martin
AI Center (LAIC)

Systems Engineering Challenges for AI

AI Risks

Policy

Human Governance

Non-determinism & Bias
Characteristics of AI Training and Inference

Model Drift
Dynamic Operating Environments

Real-World Operations
Deployments from the Edge to the Cloud

Policies and Standards
MIL-STD-882E, DoDD 3000.09, DO-178C, etc.

Accreditation for Use
Systems Engineering Artifacts Needed

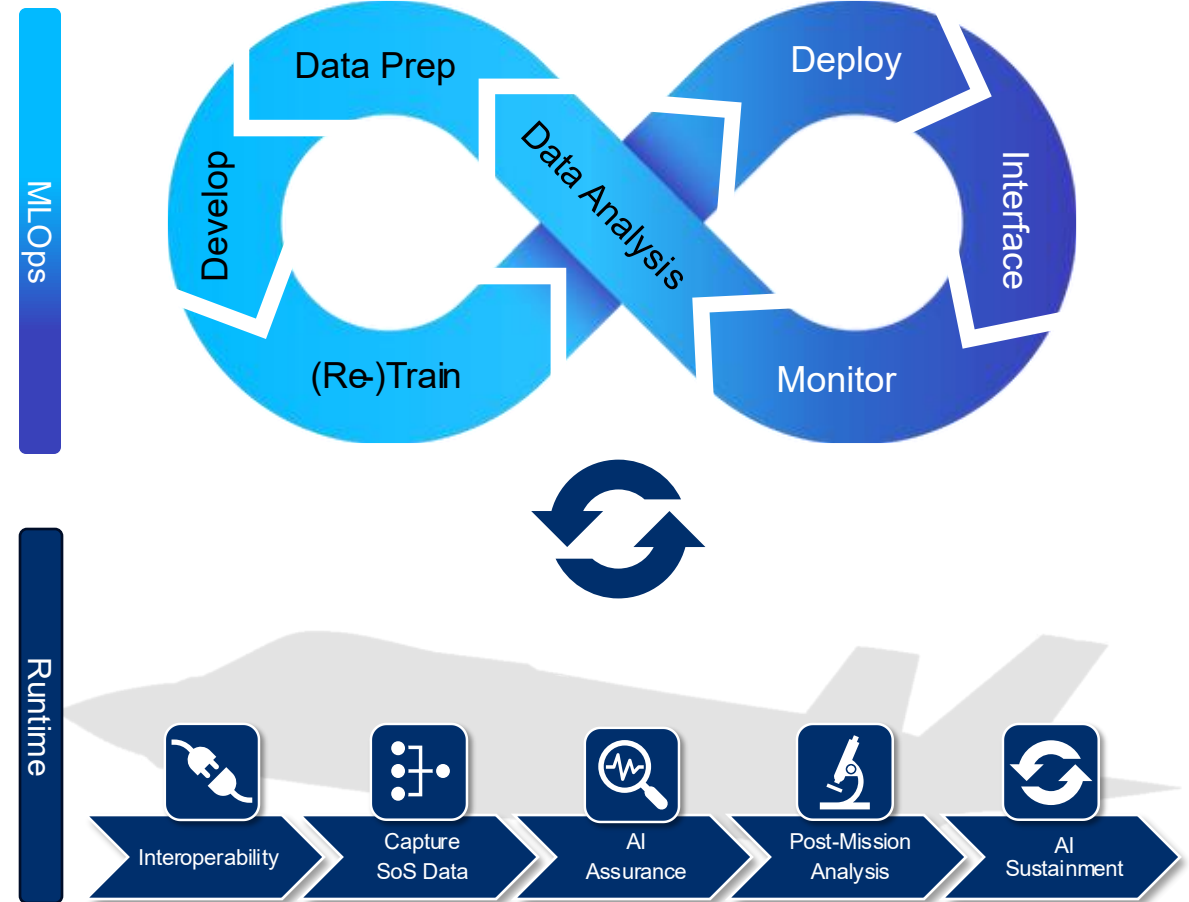**AI Governance**
Human Control and Oversight of AI

Transparency / Traceability
Help users understand and interpret AI in the processing chain
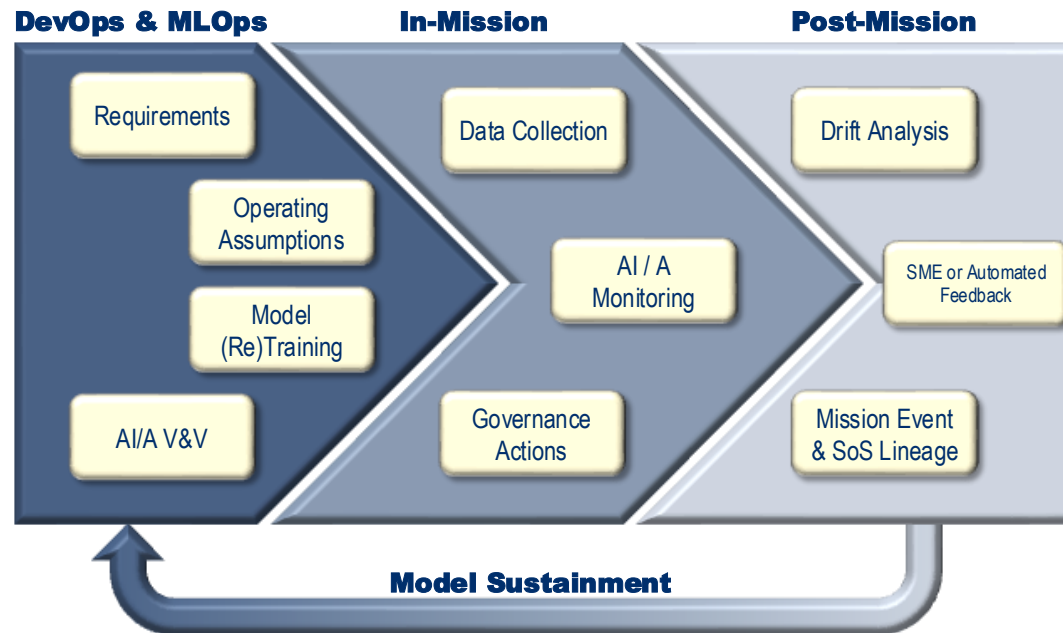
LOCKHEED MARTIN

Systems Engineering Approach

AI Assurance Framework
Open & Extensible

Capturing Mission Data
Support for AI Sustainment, Transparency for Governance & Accreditation

Support Human Roles
In Governable & Sustainable AI

Systems Eng. & Modeling
Specify information needed for AI interoperability & model governance

MLOps: Data Prep, Develop, Data Analysis, (Re-)Train, Deploy, Interface, Monitor

Runtime: Interoperability · Capture SoS Data · AI Assurance · Post-Mission Analysis · AI Sustainment

AI Model Sustainment & AI Assurance have intertwined SE Artifacts

OMEGA: AI Assurance Framework
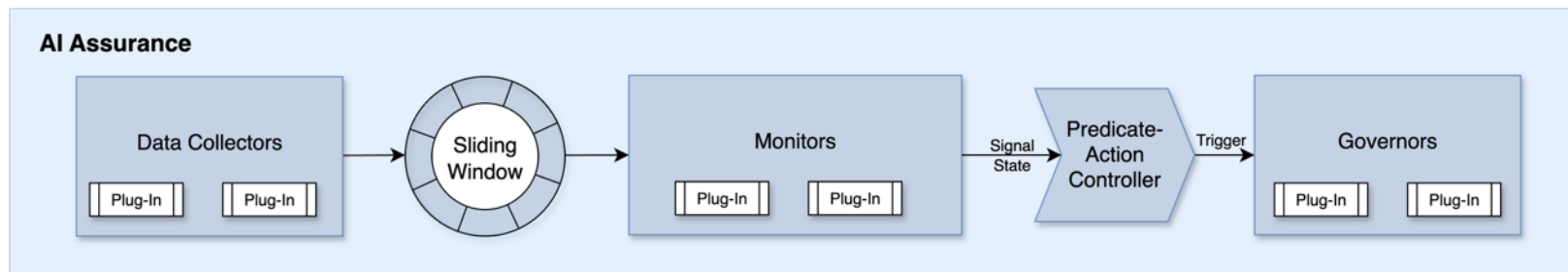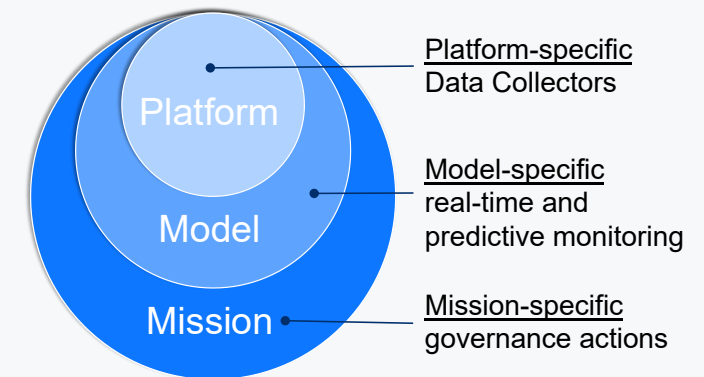## Operational Monitoring for Ethical Governance of AI / A

DevOps & MLOps

- Requirements
- Operating Assumptions
- Model (Re)Training
- AI/A V&V

In-Mission

- Data Collection
- AI / A Monitoring
- Governance Actions

Post-Mission

- Drift Analysis
- SME or Automated Feedback
- Mission Event & SoS Lineage

Model Sustainment

AI Assurance

Data Collectors → Plug-In | Plug-In → Sliding Window → Monitors → Plug-In | Plug-In → Signal State → Predicate-Action Controller → Trigger → Governors → Plug-In | Plug-In

In-Mission Monitoring
Post-Mission Analysis
Provide Data for MLOps

Is AI model performance and/or operating environment consistent with:

1. How it was Trained?
2. How it was Validated?
3. Requirements & Assumptions?

Plugin Categories:

- Platform
- Model
- Mission

Platform-specific Data Collectors

Model-specific real-time and predictive monitoring

Mission-specific governance actions

LOCKHEED MARTIN

OMEGA: Computer Vision Use Case

Predict: Car
Confidence: `0.85`
Image health: `0.95`

Predict: Boat
Confidence: `0.85`
Image health: `0.35`

Image Health "Signal"

▷ Automated Target Recognition (ATR) systems detect, classify, and identify objects of interest in a scene.

▷ Model performance depends on quality of image input defined by phenomenological degradations: blur, speckle, resolution, etc.

▷ Image degradation will determine model performance (accuracy, recall), but model may/may not incorporate input quality assessment: the model will operate on *any* input.

▷ Assessment of image health (e.g. brightness) using deterministic means provides signal to identify ATR outputs that should be flagged as unreliable: this is *separate* from confidence levels.

LOCKHEED MARTIN

Spotlight: AI / SoS Lineage

Lineage Graphs capture System of Systems Data Flow to provide context to AI Inferences.

End-User Benefits:

▶ End Users can visualize product derivation and trace its flow through.

▶ End users can provide valuable feedback on AI inferences and outputs throughout the SoS processing chain.

▶ Data, Feedback, and Context can be provided to MLOps for AI Model Sustainment.

▶ Can show "what's wrong" but also help users identify <u>why</u> it went wrong
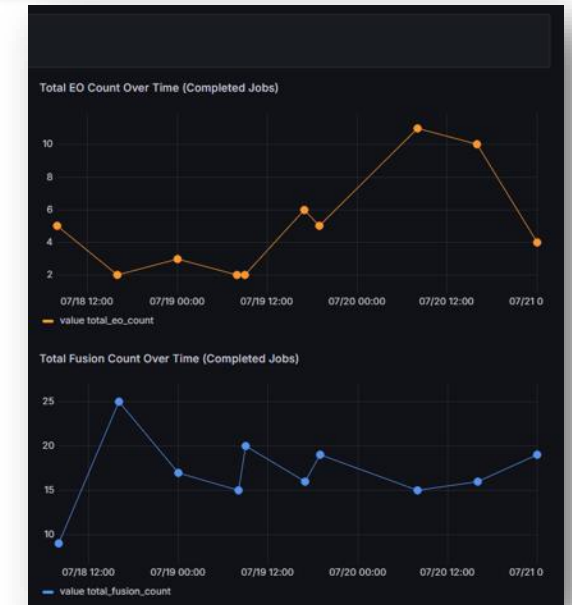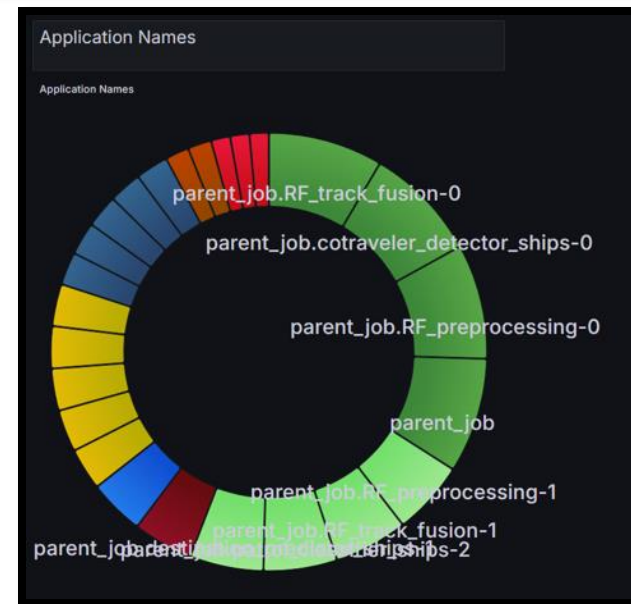
Lineage Analytics use SoS context to find and correlate insights into AI- and multi-AI Systems.

Visualizations provided by LM Space

Transparency and Context; Correlate Mission Metrics to SoS Data Flow

LOCKHEED MARTIN

MBSE DevKit & Interoperability

MBSE Start-up Kit

Resources for programs and employees starting modeling efforts

Focused Content
- MBSE articles, standards, profiles, and best practices
- Reference Architectures
- "Gold Card" information sheets
- Program planning resources
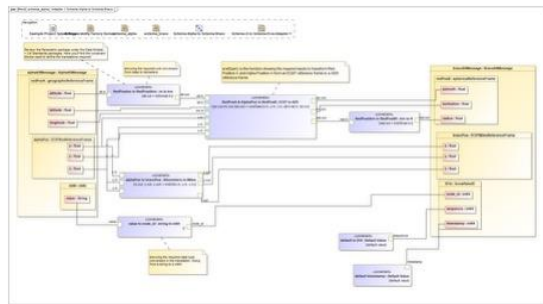- Common requirements schema
- Common Training

Complementary, integrated set of resources to support program start-up & executions

MBSE Playbook

Robust, consistent, tailorable MBSE process definition

Focused Content
- Process definition guidance (plays)
- Decision aids for programs to guide execution
- Linkage to detailed "how-to" guides and articles describing best practices

Translation Matrix

Parametric Diagram

STAR.OS™

- **Service Dev Kit (.SDK)**
- **Interoperability (.IO)**
- **User Interface (.UI)**

Systems, **T**actical Apps, **A**utonomy/AI, **R**apid Deployment

DevKit enables the STAR.OS™ Interoperability for AI-infused SoS:

- Requirements
- Interoperability
- Translators
- Configurations

The MBSE Dev Kit provides a collection of standardized tools and documentation for rapid adoption of best practices and interoperability

LOCKHEED MARTIN

Reference Architecture
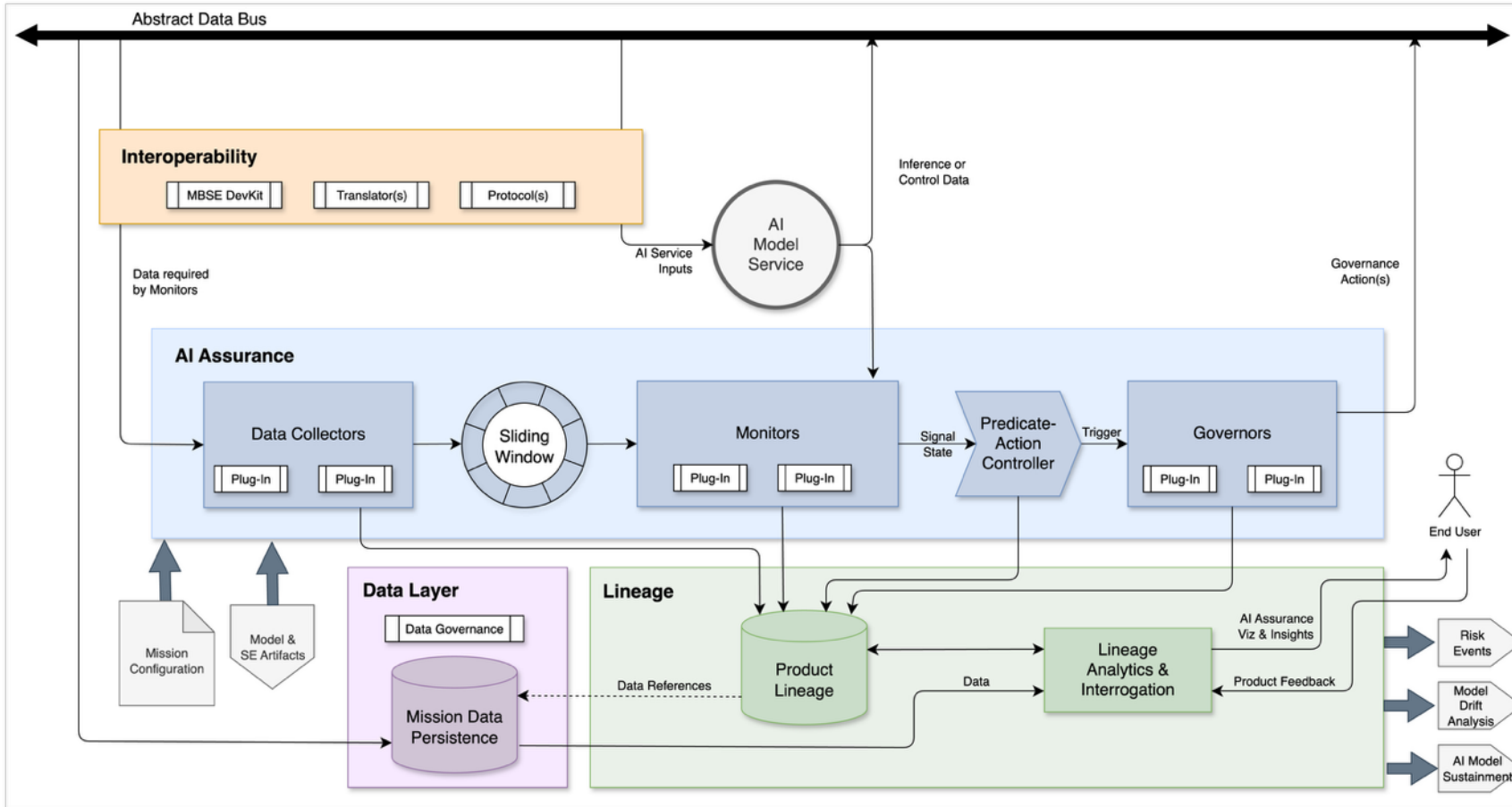
- Interoperability (STAR.OS^{TM})
  - Translators
  - Protocols / Interfaces

- AI Assurance (OMEGA)
  - AI Model Monitoring
  - Governance Actions
  - Runtime Assurance
  - Post-Mission Analysis

- Data Layer
  - Data Persistence
  - Data Governance

- Lineage (Spotlight)
  - SoS Traceability
  - Lineage Analytics
  - Lineage UX / Assistant
  - Feedback on AI Products

LOCKHEED MARTIN

Use Case: Robustness Monitoring & Runtime Assurance (Simplex)
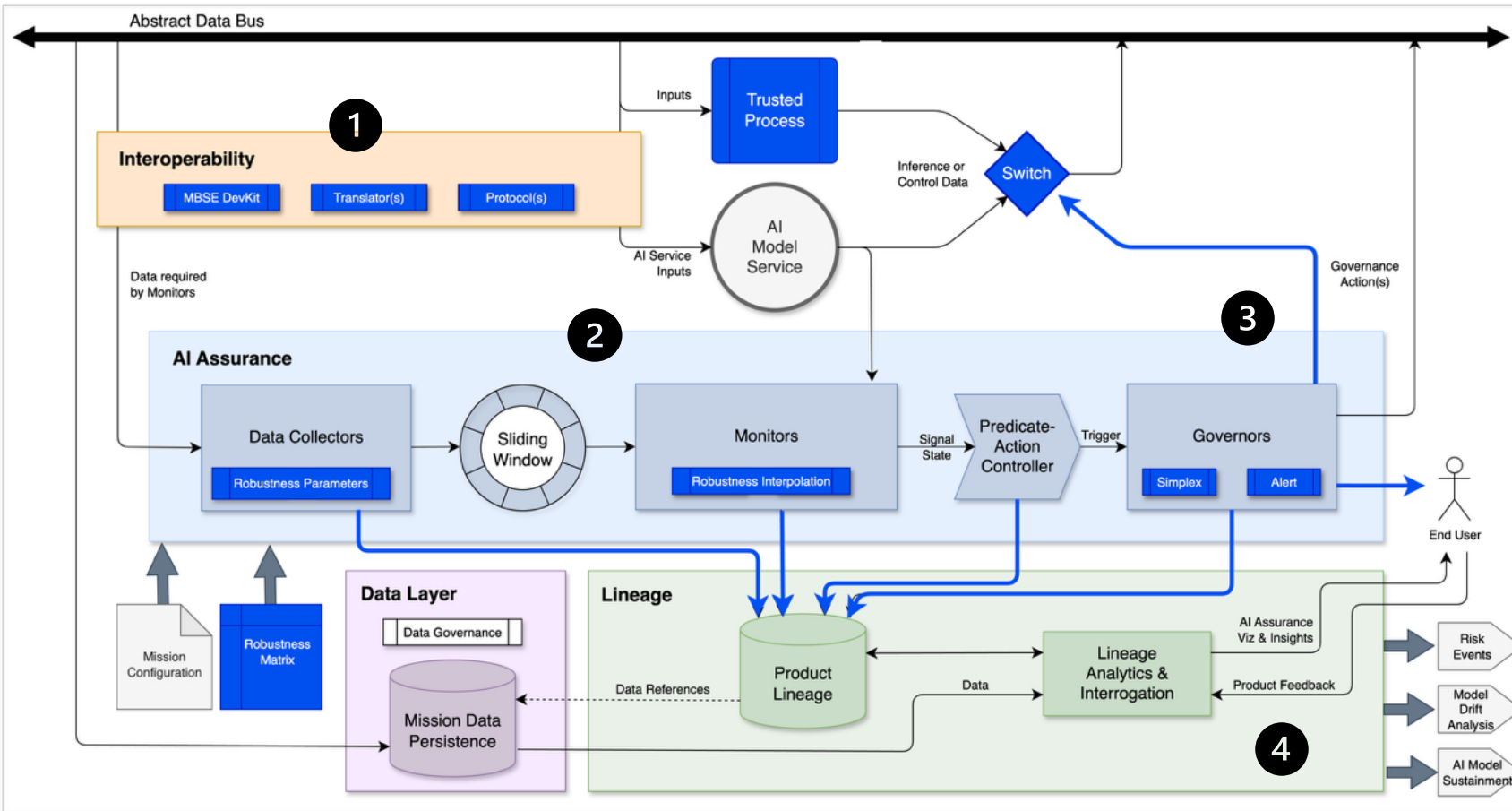
1 MBSE DevKit drives interoperability and ensures it is consistent with V&V.

2 Parametric Robustness data resulting from AI V&V can drive in-mission monitoring of AI models.
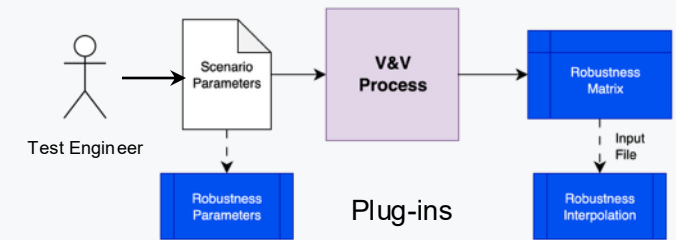
3 OMEGA can instantiate a Runtime Assurance Architecture (RTAA) Design Pattern.

4 Lineage can help explain why the RTAA was triggered and provide mission data for analysis and model re-training.

Use Case: AI V&V identifies environmental conditions that potentially result in reduced AI Model Robustness.

LOCKHEED MARTIN

System Engineering Artifacts
AI Assurance, Governance, & Sustainment, Accreditation

Left side artifacts (top to bottom):

- Human Alerting
- Explainability
- Governance Actions
- AI Feedback

- AI Assurance Criteria
- AI Lineage
- AI Sustainment Req's
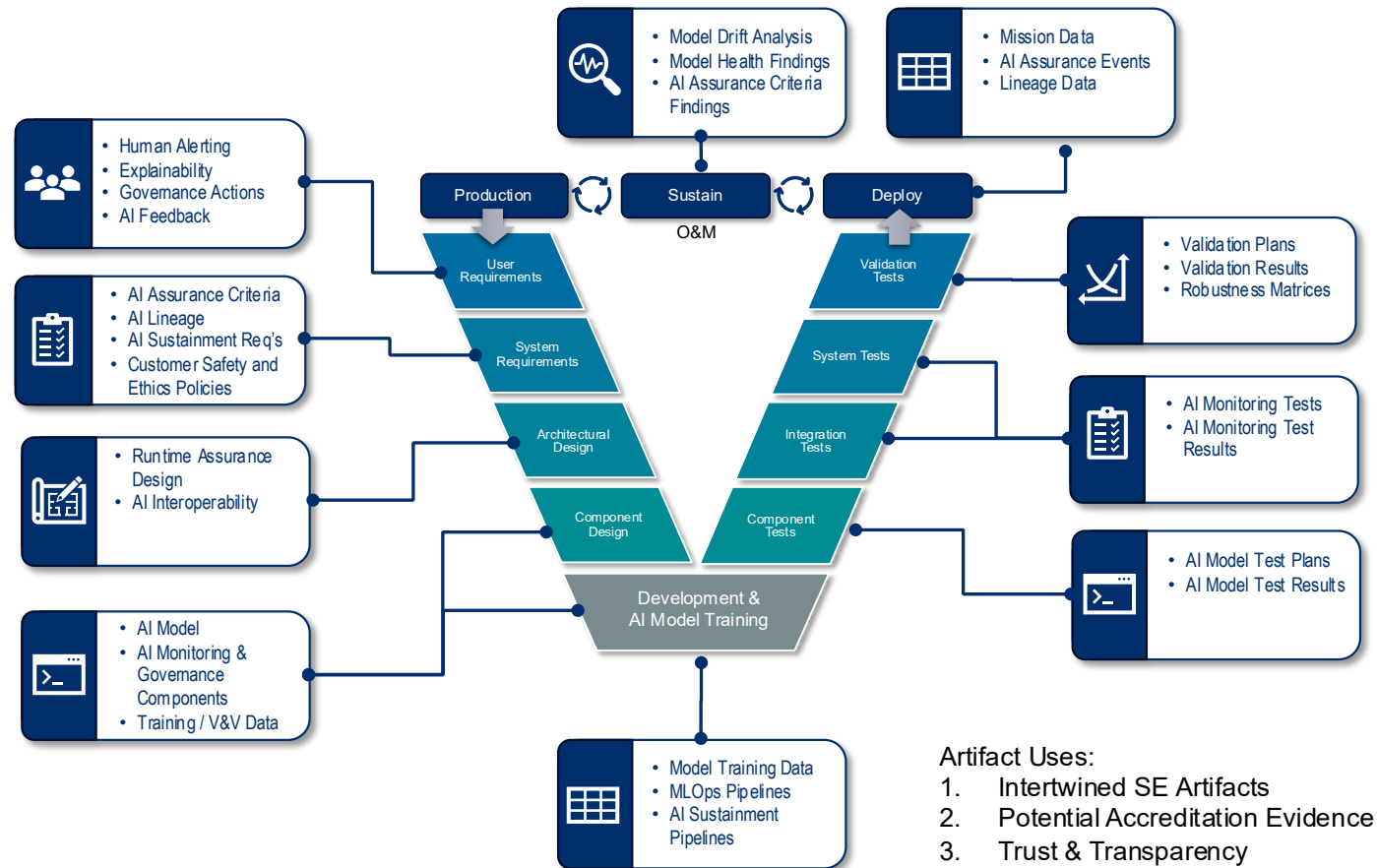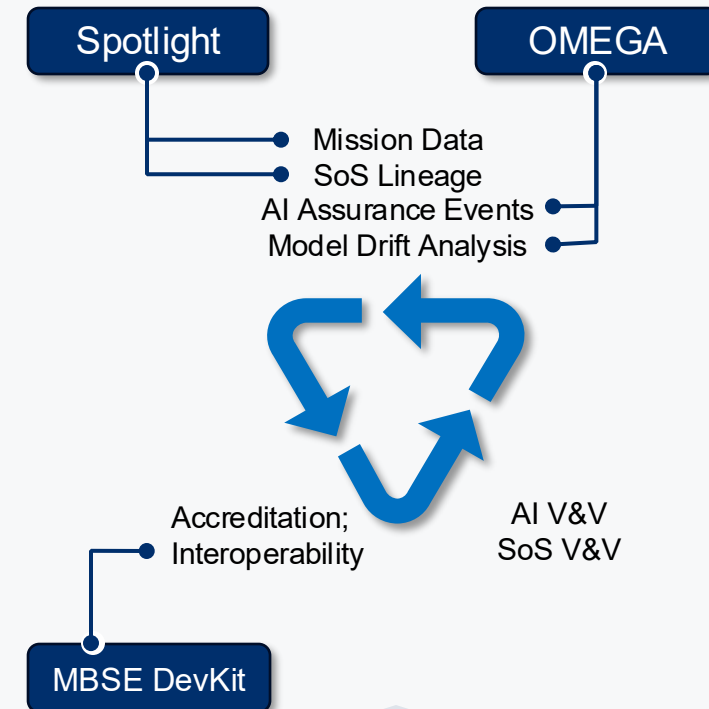- Customer Safety and Ethics Policies

- Runtime Assurance Design
- AI Interoperability

- AI Model
- AI Monitoring & Governance Components
- Training / V&V Data

Top center artifacts:

- Model Drift Analysis
- Model Health Findings
- AI Assurance Criteria Findings

- Mission Data
- AI Assurance Events
- Lineage Data

V-Model left side (top to bottom): Production, User Requirements, System Requirements, Architectural Design, Component Design

V-Model center: Sustain — O&M, Deploy

V-Model right side: Validation Tests, System Tests, Integration Tests, Component Tests

V-Model bottom: Development & AI Model Training

Right side artifacts (top to bottom):

- Validation Plans
- Validation Results
- Robustness Matrices

- AI Monitoring Tests
- AI Monitoring Test Results

- AI Model Test Plans
- AI Model Test Results

Bottom center artifact:

- Model Training Data
- MLOps Pipelines
- AI Sustainment Pipelines

Artifact Uses:
1. Intertwined SE Artifacts
2. Potential Accreditation Evidence
3. Trust & Transparency

AI Assurance & Lineage Informs MLOps

Spotlight — **OMEGA**

- Mission Data
- SoS Lineage
- AI Assurance Events
- Model Drift Analysis

Accreditation; Interoperability

AI V&V SoS V&V

MBSE DevKit

MLOps Drives AI Assurance

LOCKHEED MARTIN